

## Magyar nyelvű szövegek automatikus fonetikai átírása

Novák Attila<sup>1,2</sup>, Siklósi Borbála<sup>2</sup>

<sup>1</sup> MTA-PPKE Magyar Nyelvtechnológiai Kutatócsoport ,

<sup>2</sup> Pázmány Péter Katolikus Egyetem Információs Technológiai és Bionikai Kar  
1083 Budapest, Práter utca 50/a  
e-mail:{novak.attila, siklosi.borbala}@itk.ppke.hu

**Kivonat** Az írott szövegből hangzó beszédet előállító rendszerek egyik alapvető komponensének feladata a szöveg fonetikai átírása. Bár a betűsor–fonémásor-leképezés komplexitása nyelvenként változó, kivételek a legtöbb nyelvben vannak. Cikkünkben egy magyar szövegek fonetikai átírására alkalmas programot mutatunk be, amelyben a fonéma–graféma-átírószabályok implementálása mellett morfológiai elemző is szerepet kapott a morfémahatárok és szóösszetételek meghatározásához. Az így létrejött rendszer olyan szövegek átírása során is jól teljesít, melyek sok idegen szót tartalmaznak.

### 1. Bevezetés

A bemutatott rendszer célja írott szövegek automatikus és egyértelmű átírása azok fonetikus reprezentációjára. A rendszert egy magyar földrajzi neveket tartalmazó adatbázis átírására használtuk.

Bár az írott ábécé egységei általában megfeleltethetők a beszélt nyelv fonetikai egységeinek, ennek a leképezésnek a bonyolultsága nyelvenként változó lehet. Ha csak a latin ábécét használó nyelveket vizsgáljuk ebből a szempontból, akkor is jelentős különbségeket találunk. Ezért egy fonetikai átírórendszer mindig nyelvspecifikus, és a különböző módszerek alkalmazhatósága az adott nyelv morfoszintaktikai és fonológiai tulajdonságaitól függ.

Az angol helyesírás szabályait viszonylag korán rögzítették, azonban a kiejtési rendszer tovább fejlődött [9]. Ezért gyakran elég nehéz megjósolni az írott és kiejtett alakok közötti megfeleltetést. Mivel azonban a szóalakok száma véges, ezért akár kézzel, akár géppel készített szótár alkalmas lehet arra, hogy az angol nyelv összes szavát – azok írott és a kiejtés szerint átírt alakjaival – tároljuk. Problémát csak az új szavak és nevek megjelenése jelenthet.

Más nyelvek, például a magyar esetében az írott és kiejtett alakok általában sokkal közelebb állnak egymáshoz. A legtöbb esetben a kiejtés egyértelműen megjósolható az írott alak alapján. Mégis adódnak kivételes esetek és fonetikai korlátokból fakadó megkötések. Továbbá a magyar nyelv agglutináló jellegéből adódóan a szóalakok nem felsorolható volta miatt azok szótárban való tárolása nem lehetséges [8].

Ezért, olyan automatikus módszerre van szükség a fonetikai átírat meghatározására, ami ráadásul figyelembe veszi a technikai követelményeket, azaz a nagy mennyiségű offline adat tárolása helyett a feldolgozásra helyezi a hangsúlyt. Célnk egy ilyen rendszer megvalósítása volt.

## 2. Kapcsolódó munkák

A fonetikus átírást megvalósító módszerek három fő csoportba sorolhatók: [4]:

- szótárban való megfeleltetés,
- szabályalapú módszerek,
- adatvezérelt módszerek.

A szótárban való megfeleltetést akkor alkalmazhatjuk, ha az írott és a fonetikus reprezentáció csak konvenciókon alapul, nincsenek alkalmazható szabályszerűségek vagy általánosítások. Ezeknek a módszereknek az az előnye, hogy a szótárban kiegészítő információkat (hangsúly, szófaj) is tárolhatunk. Hátrányuk viszont, hogy a szótárak kézzel való létrehozása nagyon költséges és nehéz feladat.

Agglutináló nyelvek esetén azonban, függetlenül az agglutináció korlátozottságától, mindig vannak olyan szóalakok, amik egy előre létrehozott szótárból hiányoznak. A szabályalapú módszerek ezt a problémát előre meghatározott átírószabályok alkalmazásával oldják meg. Ezeket a nyelvspecifikus szabályokat nyelvészek definiálják, majd valamilyen keretrendszerben (pl. véges fordítóautomaták [7]) formalizálva alkalmazhatóak. Az ilyen szabályalapú megoldásoknak is szükségük van egy kivételszótárra, amiben a rendhagyó kiejtésű szóalakok átíratát rögzítik.

Gépi tanulási módszerek alkalmazására is láthatunk példát fonetikai átírást megoldására. [5] szerzői megmutatták, hogy az ilyen módszerek általánosító képessége jobb, mint a szabályalapú rendszereké (az angol nyelv esetén). Az egyik legismertebb ilyen implementáció a *Pronunciation by Analogy* (PbA), azaz a hasonló szavak átíratán alapuló átírást elvén működik [6]. Emögött az elmélet mögött az a pszicholingvisztikai modell áll, ami egy szó kiejtését a hozzá hasonló alakú, ismert kiejtésű szavak alapján határozza meg. Egy másik megközelítés, az ún. *Joint-sequence model* egy írott szóalak legvalószínűbb kiejtését a Bayes-szabály alapján határozza meg. Minden ilyen adatvezérelt megközelítés esetén szükség van azonban egy szótárra, vagy egy átírt korpuszra amiből a statisztikák meghatározhatók, illetve a gépi tanulási módszerek számára tanítóanyagként szolgál.

Magyar nyelvre létezik egy online kiejtésszótár, ami 1,5 millió szóalakot és azok fonetikai átíratát tartalmazza [2]. A szótár létrehozása több lépésből állt. Először, egy nagy írott korpusz szavait gyűjtötték össze és ebből a listából eltávolították az idegen, illetve a helyesírási hibát tartalmazó szavakat. Ezután átírószabályokat alkalmaztak. Végül, a kivételes esetek meghatározása után ezeket kézzel javították. A szótár szerzői szerint így a szótár referenciaként használható, hiszen ez magyar nyelvre a legnagyobb, IPA átíratot tartalmazó és helyesnek tekinthető szótár. Azonban, a szótár korlátja, hogy csak az eredeti korpuszban

szereplő szavakat ismeri, más szóalakok vagy új szavak átírására nem alkalmazható. Emellett az online felhasználói felületben implementált korlátozások miatt tömeges lekérdezésre nem használható.

### 3. Módszer

Fonémikus nyelvek esetén (pl. finn, észt, magyar, stb.) az írott alakok kiejtés szerinti átírása majdhogynem egyértelmű. Például az *ablak* szó kiejtése [ɒblok]. (Az 1. táblázat tartalmazza az egyes magyar betűkhöz tartozó kiejtés Nemzetközi Fonetikai Ábécé szerinti átírását.) Van azonban két olyan jelenség, ami az átírást megnehezíti: bizonyos hangkapcsolódások, illetve a régies, vagy idegen szavak. További problémát jelent a nyelv szemiotikus rendszerének (számok, rövidítések) normalizálása.

1. táblázat. A magyar betűknek megfelelő fonémák IPA jelöléssel

betű	IPA	betű	IPA	betű	IPA	betű	IPA
á	a:	b	b	n	n	zs	ʒ
a	ɔ	p	p	ny	ɲ	s	ʃ
o	o	d	d	j	j	cs	tʃ
u	u	t	t	h	h	l	l
ü	y	g	g	v	v	r	r
i	i	k	k	f	f	dz	ɖʒ
é	e:	gy	ʝ	z	z	dzs	ɖʒ
ö	ø	ty	c	sz	s		
e	ɛ	m	m	c	ts		

Rendszerünk három komponensből áll: (1) a Humor morfológiai elemzőből [10,11], (2) egy a rendhagyó szótöveket tartalmazó lexikonból és (3) a fonológiai szabályok XFST-ben (Xerox Finite-State Tool) [3] való implementációjából.

#### 3.1. Morfológiai elemzés

A feldolgozás során először a szóalakok morfológiai szerkezetét határozzuk meg, hogy a megfelelő morfémahatárokon érvénybe lépő morfofonológiai szabályok érvényességi körét megállapítsuk (pl. lexikai palatalizáció, hosszú mássalhangzók vs. szóösszetételek, stb.). Továbbá, a magyarban igen gyakori összetett szavak esetén is előfordulhatnak rendhagyó elemek, amelyek felismeréséhez az összetételi határoknál történő szegmentálásra van szükség. Ráadásul, több fonémát bigráfokkal ír le a magyar nyelv (*cs*, *gy*, *ty*, *ny*, *sz*, *zs*, *dz*, *dzs*, és ezek hosszú alakjai). De ha morfémahatáron szerepelnek egymás mellett ezek a mássalhangzók, akkor külön ejtjük őket (más szabályok miatt előfordulhat részleges-, vagy teljes hasonulás). Például az *eszközsáv* helyes átírata [eskøʃa:v], nem pedig [eskøʒa:v].

Lehetnek továbbá olyan összetett szavak, amelyeknek az egyes összetevői rendhagyó kiejtésűek. Ezek felismeréséhez is a morfológiai elemzőre van szükség annak elkerülése érdekében, hogy az általános fonológiai szabályok szerint írjuk át őket.

### 3.2. Rendhagyó szótövek

Minden nyelvben vannak rendhagyó kiejtésű szavak. Ezek általában tulajdonnevek és idegen szavak, amik azonban valamilyen mértékben idomulnak a befogadó nyelv fonológiai rendszeréhez. Ilyen példa az angol *file* szó, ami a magyar helyesírás szerint használható az eredeti formájában is, de a kiejtéshez alkalmazkodó formában is írható (*fájl*). Mindkét esetben a [fa:jl] a szó fonetikai átírata. Ezzel szemben a *New York* kifejezést csak ebben az alakban írhatjuk, kiejtése pedig [nu:jork]. Az idegen szavak mellett a hagyomány szerinti írásmóddal rendelkező szavak is hasonlóan viselkednek, így sok családnév, földrajzi név is ebbe a kategóriába tartozik.

További rendhagyó esetek az olyan szóalakok, amelyekben a helyesírási norma nem a köznyelvi kiejtésnek megfelelően jelöli valamely hang hosszúságát (pl. az *egyesület* szóalak kiejtésének az írott alak szerinti [ɛʃɛʃylɛt] helyett az [ɛʃːɛʃylɛt] felel meg).

Szintén a lexikonban vannak felsorolva az olyan rövidített, illetve különböző szimbólumokat tartalmazó kifejezések, melyek kiejtése nem az írott alak megfelelője (pl. számok, képletek, dátumok, mértékegységek, stb.). Ide sorolható továbbá a rövidítések különböző típusainak kezelése is. Ahhoz, hogy ezekhez a szóalakokhoz a kiejtés hozzárendelhető legyen, ezeket egy előfeldolgozási lépésben normalizálni kell. A nyelv szemiotikus rendszerének normalizálása számos alfeladatot tartalmaz, ezek részletes tárgyalása megtalálható [13]-ban. A rövidítések kérdéséről viszont itt is ejtünk néhány szót. A rövidített alakok kiejtése során több esettel állhatunk szemben:

- a rövidítést úgy ejtjük ki, mintha egy valódi (itt: idegen) szó lenne (pl. *NATO* [na:to:]),
- a rövidített alak helyére szóban annak kifejtett változatát helyettesítjük be (pl. *du.* [de:luta:n]),
- a rövidített alakot betűzve ejtjük ki (pl. *USB* [u:ɛʃbe:]).

Rendszerünkben a rövidítésként felismert alakokat először egy olyan listával egyeztetjük, amiben a szóként kiejtett rövidített alakok vannak felsorolva. Ha ebben nem szerepel, akkor az az alapértelmezett szabály érvényesül, hogy betűzve ejtjük ki a rövidítést.

### 3.3. Fonológiai szabályok

A harmadik komponens a helyesírás által nem jelölt fonológiai szabályok XFST-beli implementációja. A szabályok [12] leírását követik. A szabályok alkalmazásának sorrendjét több tényező határozza meg (a teljes sorrendet 1. a 2. táblázatban). A mássalhangzók helyesírási sajátosságainak kezelése megelőzi a többi

szabályt. A lexikai szabályokat a posztlexikális folyamatok végrehajtása előtt kell alkalmazni. Ezen kívül számos további jelenség egyedi kezelésére volt szükség. Ez szintén hatással volt a szabályok alkalmazási sorrendjére, ezeket lejjebb részletezzük.

2. táblázat. Fonológiai szabályok az alkalmazásuk sorrendjében

#	szabály
1.	hosszú digráfok átírása, x, w, qu, y, ly
2.	lexikai h-törlés
3.	lexikai palatalizáció
4.	lexikai palatális összeolvadás (a lex. palatalizáció megelőzi)
5.	a több szótagú tövek végén álló magas magánhangzók rövidülése (opcionális)
6.	az intervokalikus és szóvégi <i>dzs</i> és <i>dz</i> nyúlása
7.	minden szó első szótagja hangsúlyt kap
8.	zöngésségi hasonulás (regresszív, a jobb kontextust a kimeneten kell ellenőrizni)
9.	adaffrikáció (a zöngésségi hasonulás megelőzi)
10.	nazális hasonulás
11.	degemináció
12.	<i>j</i> : a fon. frázis végén: zöngétlen réshang zöngétlen obstruensek után; réshang lesz zöngés mássalhangzók után a fon. frázis végén
13.	a <i>h</i> posztlexikális váltakozása (szonoránsok után zöngéssedik; kódában palatalizáció és velarizáció)
14.	posztlexikális palatalizáció
15.	zár- és réshangok, nazálisok, likvidák: gemináció minden határon keresztül
16.	affrikáták: gemináció csak a toldalékhatárokon
17.	magánhangzók átírása

### *Helyesírási sajátosságok kezelése*

1. Ez a szabály egyrészt a digráffal jelölt palatális és szibiláns mássalhangzók (pl. *ty*, *sz*, *zs* stb.) gemináta (hosszú) alakjának kezelésére (pl. *tty*, *ssz*, *zsz* stb.) szolgál. Az ilyen alakú betűkapcsolatok valódi mássalhangzókapcsolatokat is jelölhetnek, pl. a *ssz* betűsorozat *s+sz* [Ss] kapcsolat is lehet,

ez azonban csak akkor lehetséges, ha a két mássalhangzó között morfémahatár húzódik. Emellett a szabály a csak idegen szavakban és neveken előforduló *q*, *w*, *x* és a *ty*, *gy*, *ny*, *ly* betűkapcsolatokon kívüli *y* betűk (részben kontextusfüggő) kiejtését adja meg.

### *Lexikális folyamatok*

2. A *h*-végű szavak egy részénél (pl. a *méh* szó esetében) a tővégi *h*-t általában nem ejtjük, ha nem követi magánhangzó kezdetű toldalék.
3. Az inflexiós toldalékok elején álló *j* palatalizálja az előtte álló tővégi zárhangokat és a *l*-et. Ez a lexikális palatalizációs szabály csak inflexióstoldalék-határokon működik.
4. Az inflexiós toldalékok elején álló *j* palatális geminátává olvad össze az előtte álló tővégi palatális mássalhangzókkal. A lexikai palatalizáció megelőzi (táplálja) ezt a szabályt.
5. A hosszú felső nyelvvállású magánhangzóra (*í*, *ú*, *ű*) végződő több szótagú töveket a köznyelvben általában rövid tővégi magánhangzóval ejtjük, kivéve a rendkívül igényes „színpadi” kiejtést. Ezt az opcionális rövidülést implementáltuk a rendszerben.
6. Az intervokális és a szóvégi *dzs* és *dz* ejtése hosszú. Ez alól csak néhány lexikai kivétel van, pl. a *fridzsider* [fridzider] szó.

### *Hangsúly*

7. A hangsúlykijelölés meglehetősen egyszerű a magyarban: a hangsúly mindig az első szótagra esik. Kizárólag a hangsúlytalan elemek: a determinánsok és klitikumok, illetve a tágabb tagmondat- vagy fráziskontextusban fellépő hangsúlytörlések okoznak problémát, ezeket azonban az xfst-szabályrendszeren kívül kezeltük.

### *Posztlexikális szabályok*

8. A magyarban regresszív (jobbról balra ható) zöngességi hasonulás érinti az obstruenseket. Két kivételes hang van ebből a szempontból: a *v* zöngétlenedik, de nem zöngétlenít, a *h* pedig zöngétlenít, de nem zöngésedik. Az a folyamat megelőzi az adaffrikációt.
9. Adaffrikáció: bizonyos zárhang-réshang és zárhang-affrikáta kapcsolatok a megfelelő affrikátává olvadnak össze. Nem implementáltuk a szabályrendszerben azokat az adaffrikációs jelenségeket, amelyek csak lezser, gyors beszédben érvényesülnek, mint a szóhatárokon átívelő zárhang-réshang adaffrikáció vagy a palatális-zárhang, illetve palatális-affrikáta adaffrikáció.
10. A nazális *n* képzési helye hasonul az azt követő zárhang vagy nazális képzési helyéhez, illetve az *n* és *m* labiodentális nazálisként [ŋ] realizálódik, ha labiodentális frikatíva követi.
11. Számos degeminációs folyamat van, amelyek különböző környezetekben játszódnak le. A monomorfemikus gemináták bármely más mássalhangzó környezetében degeminálódnak (rövidülnek): CC-X → C-X, X-CC → X-C (ahol

a - bármilyen morfémahatár lehet, illetve nincs is szükség morfémahatár jelenlétére). A morfémahatáron keresztül történő degemináció  $XC-C \rightarrow X-C$ ,  $C-CX \rightarrow C-X$  csak akkor kötelező, ha az  $X$  obstruens (és nazális környezetben is implementáltuk a folyamatot, mert ebben a környezetben is gyakori). A  $C-CX \rightarrow C-X$  degemináció csak az obstruensek egy részhalmazát érinti. A likvidákat követő  $LC=C \rightarrow L=C$  degemináció csak inflexióstoldalék-határokon következik be.

12. A szóvégi  $j$  zöngétlen  $[\zeta]$ , illetve zöngés frikatívaként  $[j]$  realizálódik, ha zöngés, illetve zöngétlen mássalhangzót követ.
13. A  $h$  posztlexikális váltakozást mutat. Intervokális, illetve magánhangzó és szonoráns közötti helyzetben zöngésedik. Elöl képzett magánhangzót követő kódában  $[\zeta]$ -vé palatalizálódik, egyéb esetben kódában  $[x]$ -vá velarizálódik.
14. Posztlexikális palatalizáció: a dentális  $t$ ,  $d$ ,  $n$  palatalizálódik a palatális  $ty$ ,  $gy$ ,  $ny$  előtt.
15. A zár- és réshangok, a nazálisok és a likvidák minden morfémahatáron keresztül geminálódnak.
16. A nem túl lezser beszédben az affrikáták csak a toldalékhatárokon geminálódnak.
17. Végül a hosszú magánhangzók reprezentációját is a  $V$ : jelölésre konvertáljuk.

#### 4. Kiértékelés

A rendszer kiértékeléséhez George Orwell 1984 című regényét használtuk. A sok egyedi és idegen szóösszetétel és szóalak miatt esett a választás erre a szövegre. A rendszerünk által létrehozott átiratot a több nyelvre is elérhető eSpeak rendszer [1] magyar implementációjával létrehozott átirattal hasonlítottuk össze. Az eSpeak használható úgy, hogy kimeneteként IPA átiratot kapjunk. Megjegyzendő, hogy az online elérhető<sup>3</sup> magyar elektronikus kiejtési szótár [2] használata a kiértékelés során felmerült, azonban a hozzáférési felület nem volt alkalmas arra, hogy automatikus kiértékelésre használjuk. Ebben a szótárban 1,5 millió szóalak található, ragozott alakokkal együtt, ami jól reprezentálja a magyar nyelvet és 99%-ban helyes. Azonban az adatbázis nem letölthető és a leírásában szereplő első 1000 találatot elmentő funkció sem elérhető, ezért nem tudtuk felhasználni.

A valódi összehasonlításhoz az eSpeak rendszert is kiegészítettük, mert a posztlexikális hasonulások (zöngésségi hasonulás, palatalizáció, nazális hasonulás, a  $/h/$  és  $/j/$  hasonulása) nincsenek jelölve az eSpeak kimenetén, ezen kívül nem különbözteti meg az affrikátákat a zárhang-réshang kapcsolatoktól (pl.  $/tj/$  vs.  $/tʃ/$ ), illetve a gemináta mássalhangzók reprezentációja gyakran hibás (pl.  $/t:/$  helyett  $/tt/$ ). Ezeket a hibákat a kimenet utófeldolgozásával javítottuk, hogy helyes és a saját rendszerünkkel összehasonlítható átiratok keletkezzenek. Egy másik eltérés a két rendszer között a tővégi hosszú felső nyelvállású magánhangzók opcionális rövidülése, amit a saját rendszerünkben implementáltunk, az eSpeak viszont ezeket a kicsit mesterkélt finomkodó hosszú alakjukban írja át. Ezeknek a hangoknak a rövid kiejtése jellemzőbb a köznyelvi magyarban.

<sup>3</sup> <http://beszedmuhely.tmit.bme.hu/mksz/>

A kiértékelés során hibaarányt mértünk (szó szinten) a teljes korpuszon vizsgálva. Azokban az esetekben, ahol több helyes átírat is helyes, bármelyik változatot elfogadtuk. A két rendszer eredményeit a 3. táblázat tartalmazza. Az eredeti eSpeak rendszer szóhibaaránya (WER) 14,81%, a kiegészített eSpeak hibaaránya 0,98%, míg a saját rendszerünk által elért hibaarány csupán 0,35% volt. Látható tehát, hogy a tesztszövegben előforduló idegen, illetve kitalált szavak sem okoztak gondot az átíróprogram számára.

3. táblázat. Kiértékelés. u/i: a rövidülő tövégi hosszú felső nyelvállású magánhangzókat tartalmazó szavak aránya; hason/h/j/N/zöng: azon szavak aránya, amelyekben a zöngésségi/palatális/nazális/j/h-hasonulás hibásan nincs jelölve, de ettől eltekintve helyesek; WER: a maradék szóhiba-arány.

rendszer	WER
a mi rendszerünk WER	0,35%
eSpeak u/i	0,98%
eSpeak WER	2,26%
eSpeak hason/h/j/N/zöng	14,81%

Az eSpeak kimenetében tapasztalt korábban nem említett hibák elsősorban a következő okokra vezethetők vissza: Lexikai hiányok (ide értve a szövegben szereplő számos angol név kiejtését), gyakori rövidítések nem megfelelő feloldása, a gemináta /r/ és a *ch* digráf kiejtésével kapcsolatos hibák, néhány szó kiejtésének ábrázolásával kapcsolatos idioszinkratikus hibák, és a lexikai palatalizáció túllakalmazása olyan helyeken, ahol nem lenne szabad megjelennie. Az utóbi hibát a morfológiai elemzés hiánya okozza: a lexikai palatalizációt mintaillesztési módszerrel kezelik az eSpeakben, és a minta ott is illeszkedik, ahol nem kéne.

Az általunk implementált rendszernek sokkal jobban megy az angol nevek kiejtése, hibáit elsősorban (az eSpeakétől különböző) lexikai hiányok, egyes rövidítések hibás feloldása, és egyes állószóösszetételek túlélemzése okozza. Az Orwell által kreált az 1984-ben szereplő újbeszél szavak egyik rendszernek sem okoztak komoly fejtörést, mert a kiejtésük szabályos, és mindkét rendszer algoritmikus átírókomponenst tartalmaz ahelyett, hogy pusztán szótárra támaszkodna.

## 5. Konklúzió

Bemutattunk egy magyar nyelvű szövegek automatikus fonetikai átírására alkalmas automatikus eszközt. A rendszer nem csak egyes szavakat képez le azok egy szótárban található átíratára, hanem teljes mondatok átírására is alkalmas, mivel figyelembe veszi a szóhatárokon előforduló hasonulásokat. Ezt egy, a morféma-, és



összetételi határok meghatározására képes morfológiai elemző és fonetikai átíró-szabályok alkalmazásával valósítottuk meg. Bemutattuk továbbá, hogy nagyméretű lexikon nélkül is jó minőségű fonetikai átírás állítható elő, hiszen a rendszer nem korlátozódik egy előre létrehozott lexikonban eltárolt szavak kezelésére. Ez a funkció egy olyan nyelv esetén, mint a magyar, ahol újabb és újabb szóalakok fordulhatnak elő, kiemelkedő fontosságú. Megmutattuk, hogy egy sok idegen szót tartalmazó korpuszon való kiértékelés során a rendszerünk jóval alacsonyabb hibaaarányal teljesít, mint egy kereskedelmi eszköz, aminek a kimenetét ráadásul sokkal kevésbé szigorúan kezeltük.

## Hivatkozások

1. eSpeak. <http://espeak.sourceforge.net/>, accessed: 2015-04-10
2. Abari, K., Olaszy, G., Zainkó, Cs., Kiss, G.: Magyar kiejtési szótár az interneten. In: IV. Magyar Számítógépes Nyelvészeti Konferencia. pp. 223–230. SZTE, Szeged (2006)
3. Beesley, K., Karttunen, L.: Finite State Morphology. No. 1 in CSLI studies in computational linguistics: Center for the Study of Language and Information, CSLI Publications (2003), <http://books.google.hu/books?id=59RoAAAAIAAJ>
4. Bisani, M., Ney, H.: Joint-sequence models for grapheme-to-phoneme conversion. *Speech Commun.* 50(5), 434–451 (May 2008), <http://dx.doi.org/10.1016/j.specom.2008.01.002>
5. Damper, R., Marchand, Y., Adamson, M., Gustafson, K.: Evaluating the pronunciation component of text-to-speech systems for english: a performance comparison of different approaches. *Computer Speech and Language* 13(2), 155 – 176 (1999), <http://www.sciencedirect.com/science/article/pii/S0885230898901176>
6. Dedina, M.J., Nusbaum, H.C.: Pronounce: a program for pronunciation by analogy. *Computer Speech and Language* 5(1), 55 – 64 (1991), <http://www.sciencedirect.com/science/article/pii/088523089190017K>
7. Kaplan, R.M., Kay, M.: Regular models of phonological rule systems. *Comput. Linguist.* 20(3), 331–378 (Sep 1994), <http://dl.acm.org/citation.cfm?id=204915.204917>
8. Kurimo, M., Puurula, A., Arisoy, E., Siivola, V., Hirsimäki, T., Pytkönen, J., Alumäe, T., Saraclar, M.: Unlimited vocabulary speech recognition for agglutinative languages. In: Proceedings of the Main Conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics. pp. 487–494. HLT-NAACL '06, Association for Computational Linguistics, Stroudsburg, PA, USA (2006), <http://dx.doi.org/10.3115/1220835.1220897>
9. Németh, G., Olaszy, G.: A magyar beszéd. Akadémiai Kiadó, Budapest, Hungary (2010)
10. Novák, A.: Milyen a jó Humor? In: I. Magyar Számítógépes Nyelvészeti Konferencia. pp. 138–144. SZTE, Szeged (2003)
11. Prószéky, G., Kis, B.: A unification-based approach to morpho-syntactic parsing of agglutinative and other (highly) inflectional languages. In: Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics. pp. 261–268. ACL '99, Association for Computational Linguistics, Stroudsburg, PA, USA (1999)

12. Siptár, P.: A magánhangzók. In: Kiefer, F., Bánréti, Z., Ács, P. (eds.) Fonetológia. No. 2 in Strukturális magyar nyelvtan, Akadémiai Kiadó (1994), <http://books.google.hu/books?id=j6xiAAAAMAAJ>
13. Taylor, P.A.: Text-to-speech synthesis. Cambridge University Press, Cambridge, UK, New York (2009), <http://opac.inria.fr/record=b1129276>